

# M/M/1 queue

Nicola Marchetti

[nicola.marchetti@tcd.ie](mailto:nicola.marchetti@tcd.ie)

## Deriving queuing theory results

- Poisson assumptions for the arrival of customers to a queue make the analysis tractable
  - Markov chains
- Useful references:
  - T. G. Robertazzi, *Computer Networks and Systems*, 3<sup>rd</sup> edition, Springer, 2000
  - L. Kleinrock, *Computer Applications, Volume 2, Queuing Systems*, Wiley, 1976

## Little's law

- The expected number in the system (queue) is the product of the arrival rate and the average time spent in the system (queue)

$$E[n] = \lambda E[\tau]$$

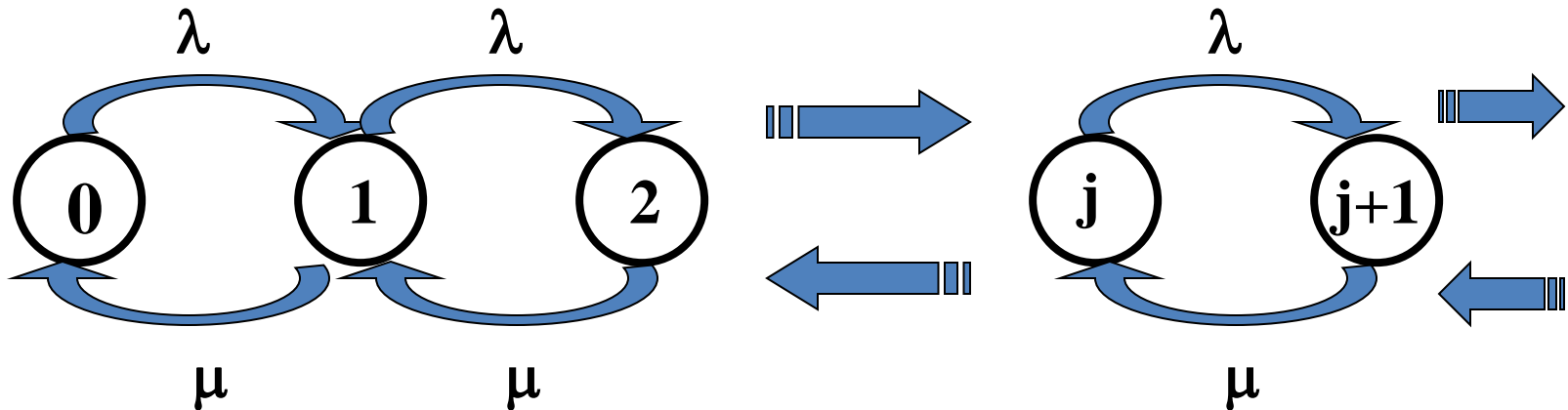
## The M/M/1 model

- Single-server system
- Arrivals according to Poisson process of rate  $\lambda$
- Inter-arrival times exponential r.v.'s
- Service times exponential r.v.'s with mean  $1/\mu$
- Infinite buffers



## Markov chain

- Due to the **memoryless property** of exponential r.v.'s, the number of customers in the system can be expressed as a continuous-time Markov chain
  - States represent the number of customers (packets, users) in the queuing system



## Probability mass function

- Let  $\rho = \lambda / \mu < 1$  (called **utilization**)
- We can then determine the probability mass function for the number of customers currently in the system
  - Let  $p_n$  denote the probability that  $n$  customers are currently in the system
- Results:

$$p_0 = (1 - \rho)$$

$$p_n = \rho^n p_0 = \rho^n (1 - \rho)$$

## Average number in the system

- We can calculate the **average number** of customers (packets, calls,...) in the system as

$$E[n] = \sum_{j=0}^{\infty} j p_j = \frac{\rho}{1-\rho}$$

## Average delay

- We can use Little's law to get the **average delay** through the system:

$$E[\tau] = \frac{1}{\lambda} \frac{\rho}{1 - \rho} = \frac{1}{\mu - \lambda}$$

- Also, the average waiting time is:

$$E[\tau_Q] = E[\tau] - E[s] = E[\tau] - \frac{1}{\mu} = \frac{1}{\mu} \frac{\rho}{1 - \rho}$$

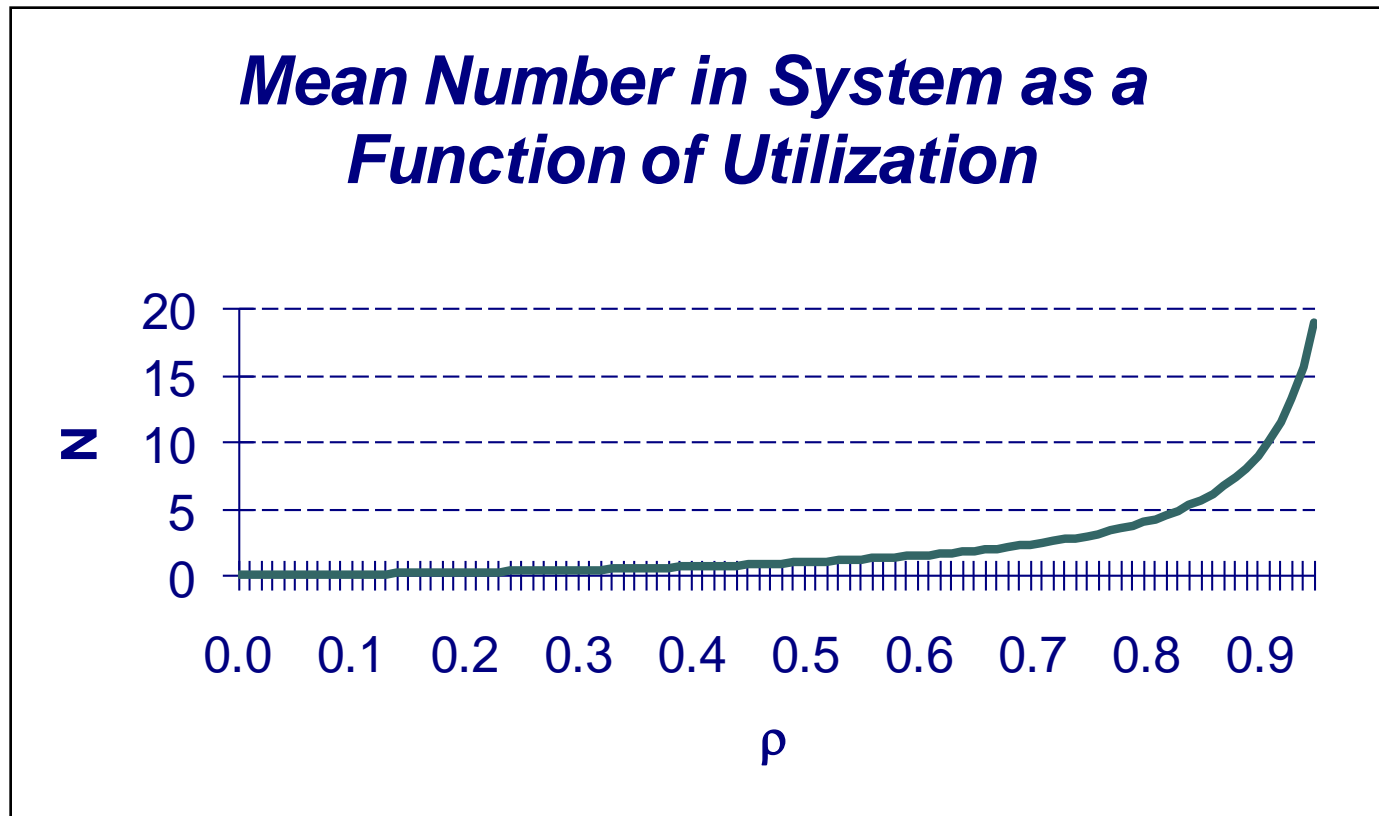


## Average number in the queue

- We can use Little's law again to get the average number of customers in the queue:

$$E[n_q] = \lambda E[\tau_Q] = \frac{\rho^2}{1 - \rho}$$

## M/M/1 queue graphical results



# Problem



A concentrator receives messages from a group of terminals and transmits them over a single transmission line. Suppose that messages arrive according to a Poisson process at a rate of one message every 4 ms, and that message transmission times are exponentially distributed with mean 3 ms.

What is the utilization of the system ?

What is the average delay through the system, expressed in ms?

- **M/M/1**
- **arrival rate**  $\lambda = 1/4$  msg/msec
- **service rate**  $\mu = 1/3$  msg/msec



What is the utilization of the system ?

$$\rho = \frac{\lambda}{\mu} = 3/4$$

What is the average delay through the system, expressed in ms?

$$E[\tau] = \frac{1}{\mu - \lambda} = \frac{1}{1/3 - 1/4} = 12$$

# Problem



You have designed a first-come first-served data processing system. You based the design on an  $M/M/1$  queuing model. The log of the system shows that there are on average 19 jobs in the system and that the average processing time (service time) for one job is 2 seconds.

Estimate the system utilization, the job arrival intensity and the average waiting time.



$$E[n] = \frac{\rho}{1 - \rho}$$

$$19 = \frac{\rho}{1 - \rho} \Rightarrow 19(1 - \rho) = \rho \Rightarrow \rho = \frac{19}{20} = 0.95$$

$$\lambda = \rho\mu = 0.95 \times \frac{1}{2} = 0.475 \text{ jobs/sec}$$

$$E[\tau_Q] = \frac{1}{\mu} \frac{\rho}{1 - \rho} = \frac{1}{1/2} \frac{0.95}{1 - 0.95} = 38 \text{ sec}$$

# Problem



In a single server system, arrivals can be modeled as a Poisson process with rate of  $3 \text{ min}^{-1}$  and the service times are exponentially distributed with mean 10 s.

Give the Kendall's notation (motivating your choice) and calculate:

- a) the average number of customers in the system;
- b) the average delay and the average waiting time;
- c) the average number of customers in the queue.





Kendall's notation: M / M / 1

Poisson  
arrivals

Exponential  
service  
times

One server

$$\rho = \frac{\lambda}{\mu} = \frac{3 / 60}{1 / 10} = 0.5$$

a)  $E[n] = \frac{\rho}{1 - \rho} = 1$

b)  $\left\{ \begin{array}{l} E[\tau] = \frac{1}{\lambda} E[n] = 20 \text{ s} \\ E[\tau_q] = E[\tau] - \frac{1}{\mu} = 10 \text{ s} \end{array} \right.$

c)  $E[n_q] = \lambda E[\tau_q] = 0.5$